

Dynamics and energetics for a molecular zipper model under external driving

Viktor Holubec¹, Petr Chvosta¹, Philipp Maass²

¹ Department of Macromolecular Physics, Faculty of Mathematics and Physics, Charles University, CZ-180 00 Praha, Czech Republic

² Fachbereich Physik, Universität Osnabrück, Barbarastraße 7, 49076 Osnabrück, Germany

E-mail: viktor.holubec@gmail.com

Abstract. We investigate the dynamics of a single-ended N -state molecular zipper based on a model originally proposed by Kittel. The molecule is driven unidirectionally towards the completely unzipped state with increasing time t . The driving lowers the energies of states with k unzipped links by an amount proportional to kt . We solve the Pauli rate equation for the state probabilities and the partial differential equations, which yield the probability distributions for the work performed on the zipper and for the heat exchanged with the thermal reservoir. Similarly to the related equilibrium model, two different regimes can be identified at a given temperature with respect to released molecular degrees of freedom per broken bond. In these two regimes the time evolution of the state probabilities as well as of the work and heat distributions show a qualitatively different behavior.

PACS numbers: 02.50.Ey, 05.70.Ce, 05.70.Ln, 05.40.-a, 82.37.-j

1. Introduction

The Watson-Crick double-stranded form represents the thermodynamically stable state of DNA in a wide range of temperature and salt conditions. However, even at standard physiologic conditions, there always exists possibility that the double-helix is locally unzipped into two strands both at its ends and in its interior [1, 2, 3, 4, 5]. If the interior unzipping is neglected, the unfolding of the two strands can be described by a simple zipper model [6, 7, 8]. This model, while including several simplifications, emphasizes the essential ingredient of the unzipping process, that means the competition between the entropic forces which tend to open the macromolecule, and the energetic forces that tend to condense it into its double-stranded form.

In the last two decades, new experimental techniques have been developed for detailed analysis of unzipping processes. A powerful technique are single molecule manipulations by optical tweezers [9, 10, 11, 12, 13, 14], where biomolecules are unfolded and refolded by applying mechanical forces. Because single molecules are subject to strong fluctuations, a stochastic description of the unfolding/refolding kinetics becomes necessary. Investigation of these stochastic processes can be useful to understand how biomolecules unfold and fold under locally applied forces [15, 16, 17, 18], as, for example, when mRNA passes through the ribosome during the translation process [19, 20], or when DNA is unzipped by helicase during the replication process [21, 22].

A particularly interesting part in the analysis of unzipping processes is the application of integral or detailed fluctuation theorems [23, 24, 25, 14, 26] to estimate free energy differences between folded and unfolded states (see, e.g., [27]). Their advantage is that they can be applied also to protocols, which drive the considered system far from equilibrium. It is thus not necessary to perform the unfolding/folding under quasi-static near-equilibrium conditions, where the process becomes reversible. Most popular theorems are the Crooks fluctuation theorem [28] and Jarzynski equality [29]. These relate to the distribution of work performed on the molecule during the process. Unfortunately it is not easy to get theoretical insight into characteristics of the underlying work distributions in far-from-equilibrium processes, since the work is a functional of the whole stochastic trajectory. Investigations have been conducted for simple spin systems driven by a time-dependent external field [30, 31, 32, 33, 34, 35] and for diffusion processes in the presence of a time-dependent potential [36, 37]. Analytical solutions are known for two-state systems [32, 33, 38, 35] and for systems with one continuous state variable [36, 37].

In this study we present analytical results for the work distribution for a multi-state system, which is motivated by a model proposed by Kittel for describing the melting transition of DNA molecules [8]. In order to derive analytical results, we had to consider a stochastic process with directed forward and forbidden backward transitions between states. As a consequence, detailed balance is broken and fluctuation theorems, as mentioned above, will not hold true. On the other hand, kinetic Monte Carlo simulations of the stochastic process show that, for at least certain parameter settings in the model,

the restriction of forbidden backward transitions is not so severe and events dominating the integrand in the Jarzynski equality are not very rare. It is important to point out that these parameter settings are not related to any experimental conditions. In more realistic settings rare events would play the decisive role and in such cases it becomes difficult to determine tails of the work distribution with sufficient accuracy. Also the relation of the work considered in our study to the thermodynamic work measured in an unzipping experiments needs to be treated with care. These problems are discussed in detail in Sec. 2 and they imply that the theory cannot be applied to experiments at this stage. Our findings should nevertheless be useful because connection to experiments seems not completely out of reach and because they widen the range of hopping models, where analytical results for work distributions are available.

2. Unzipping in an extended Kittel model

Kittel's model [8] is a simplified version of the Poland-Scheraga model [1] for the equilibrium properties of DNA molecules, which got renewed attention in the last ten years [39]. In contrast to the Poland-Scheraga model, it disregards the possibility of any interior openings (bubbles) of double-stranded parts of the molecule. Despite this simplification, it is already sufficient to understand the origin of a melting transition.

A molecule in Kittel's model [8] has N links in its fully folded state, see FIG. 1. Different states $k = 1, \dots, N$ of the molecule refer to configurations, where $(k - 1)$ of the links are opened in a row. The difference $E_k - E_1$ of the internal energy of a state with $(k - 1)$ open links and the ground state $k = 1$ (no open links) is equal to $(k - 1)\Delta$, corresponding to a loss of chemical bond energy Δ per broken link. With each broken link, the molecule gains G degrees of freedom, which characterize the win of conformational degrees of freedom when double-stranded DNA is transferred to single stranded DNA. The entropy difference $S_k - S_1$ between the state k and the ground state then becomes $k_B \ln G^{k-1} = k_B(k - 1) \ln G$, where k_B is the Boltzmann constant. The free energy F_k of state k at a temperature T is thus given by $F_k = F_0 + (k - 1)[\Delta - k_B T \ln G]$ and the equilibrium properties can be readily worked out by considering the partition sum $Z = \sum_{k=1}^N \exp(-\beta F_k)$, where $\beta = 1/k_B T$.

In extending this model to treat unfolding kinetics, the molecule is supposed to unzip successively, one link in each step, as a consequence of a strong external driving. Such external driving can be caused by a local force or a pH gradient, as indicated in FIG. 1. With respect to pulling experiments on single DNA molecules out of equilibrium, there is evidence that a neglect of interior openings can become even less relevant than

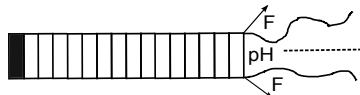


Figure 1. The zipper model.

for the equilibrium properties. For example, far away from the melting temperature, e.g. at standard room temperature conditions (298 K), the unfolding kinetics of DNA hairpin molecules could be successfully described by assuming no interior openings and thermally activated unzipping transitions [16, 38]. Moreover, molecular fraying can be identified in individual unzipping trajectories by considering the size of force jumps. Thereby states with interior openings can be systematically excluded from the analysis [40].

We assume that the driving lowers the energy differences by an amount proportional to $(k - 1)t$, i.e. open states with a larger number of broken links are favored with increasing time. If we take the ground state energy as the reference point, $E_1 = 0$, we obtain for the free energies $F_k(t) = E_k(t) - k_B T \ln G^{k-1}$ of the states

$$F_k(t) = (k - 1)[\Delta - vt - k_B T \ln G], \quad k = 1, \dots, N, \quad (1)$$

where the parameter v has the dimension of an energy rate and characterizes different speeds of the unfolding in response to different strengths of the external driving.

Following established theoretical descriptions for the rupture kinetics of the bonds [38, 40, 41], the Kramers-Bell form [42] is used for the time-dependent transition rate $\lambda_{k,k+1}$ from state k to state $(k + 1)$,

$$\lambda_{k,k+1}(t) = \tilde{\nu} \exp[-\beta F_{k,k+1}(t)]. \quad (2)$$

Here $\tilde{\nu}$ is an attempt frequency and $F_{k,k+1}(t)$ denotes the free energy barrier for the transition, i.e. the difference of the free energy at the saddle point separating states k , $(k + 1)$ and the free energy F_k in state k . The barrier $F_{k,k+1}(t)$ is considered to be composed of a bare, k -independent free energy barrier, F_b , which is modified by an amount proportional to the free energy difference, $[F_{k+1}(t) - F_k(t)]$,

$$F_{k,k+1}(t) = F_b + \gamma[F_{k+1}(t) - F_k(t)] = F_b + \gamma[\Delta - vt - k_B T \ln G]. \quad (3)$$

In the following we set $\gamma = 1$ [43]. Note that $F_{k,k+1}(t)$ and thus $\lambda(t) \equiv \lambda_{k,k+1}(t)$ in EQ. (2) are independent of k . The attempt frequency $\tilde{\nu}$ in EQ. (2) was reported [44] to be approximately proportional to the ratio of the diffusion constant of the molecule in water to the water viscosity. We assume here a linear dependence of this ratio on the temperature T , i.e. we take $\tilde{\nu} = \nu(T/T_0)$, where ν and T_0 are positive constants. Thus we obtain

$$\lambda(t) = \lambda_{k,k+1}(t) = g \exp[-\beta(d - vt)], \quad (4)$$

where

$$g = \nu G \frac{T}{T_0}, \quad d = \Delta + F_b. \quad (5)$$

If detailed balance is obeyed, i.e. $\lambda_{k,k+1}(t) \exp[-\beta F_k(t)] = \lambda_{k+1,k}(t) \exp[-\beta F_{k+1}(t)]$, the rates $\lambda_{k+1,k}(t) \equiv \lambda_b$ for backward (refolding) transitions become both independent of k and t ,

$$\lambda_b = \lambda_{k+1,k}(t) = \frac{g}{G} \exp(-\beta F_b). \quad (6)$$

Note that G appears here in the denominator, which means that the ratio $\lambda_b/\lambda(t) = \exp(\beta\Delta - \beta vt)/G$ of backward to forward rates becomes small for large degeneracy factors G . This reflects the fact that it is difficult for the flexible unfolded part of the molecule to find proper configurations, which would allow for a reformation of (hydrogen) bonds.

If the model would refer to the hopping motion of a particle between time-dependent energy levels $E_k(t)$, the work performed on the system (for one realization of the stochastic process) would be

$$\begin{aligned} W(t) &= \int_0^t dt' \sum_{k=1}^N \dot{E}_k(t') \delta_{kD(t')} \\ &= [E_{D(t)}(t) - E_{D(t)}(t_{D(t)})] + \sum_{k=1}^{D(t)-1} [E_k(t_{k+1}) - E_k(t_k)] \\ &= [F_{D(t)}(t) - F_{D(t)}(t_{D(t)})] + \sum_{k=1}^{D(t)-1} [F_k(t_{k+1}) - F_k(t_k)] , \end{aligned} \quad (7)$$

where $D(t)$ denotes the (random) state of the molecule at time t , $\dot{E}_k(t) = dE_k(t)/dt$, and t_k is the (random) time at which the transition from state k to $(k+1)$ (rupture of k th bond) takes place ($t_1 = 0$ being the initial time). The last line in Eq. (7) follows from the fact that the entropy in the extended Kittel model is not dependent on time, i.e., for given k , differences between internal energies and between free energies at distinct times are equal.

The work in Eq. (7) can be related to the thermodynamic work W_f in an unzipping experiment under force control. As pointed out in [45], work in thermodynamics is the internal energy transferred to a system upon changing the control parameters for given system configuration. This means that, when the force $f = f(t)$ is the control variable and the molecular extension x the conjugate configurational variable, one has $W_f(t) = -\int_{f(0)}^{f(t)} df x$ [46]. Stochastic changes of the molecular extension occur mainly due to bond rupture, while in between transitions the response will be rather smooth and, under neglect of small thermal fluctuations, can be represented by a function $x(k, f)$. This specifies the *mean* end-to-end distance of the unfolded part in state k at force f (as often modeled by, e.g., the freely jointed or worm-like chain models from polymer physics). With $f_k \equiv f(t_k)$ we then have

$$W_f(t) = -\int_{f(0)}^{f(t)} df x = -\int_{f_{D(t)}}^{f(t)} df x(D(t), f) - \sum_{k=1}^{D(t)-1} \int_{f_k}^{f_{k+1}} df x(k, f) . \quad (8)$$

In more detailed energy landscape models, the free energy $F_{\text{tot}}(k, f)$ of the molecule in state k under loading f can be represented as $F_{\text{tot}}(k, f) = F_0(k) - fx(k, f) + F_{\text{str}}(k, f)$ (see, e.g., [16]), where $F_0(k)$ is the free energy in the absence of loading and $F_{\text{str}}(k, f)$ the elastic energy of the unfolded part upon stretching. For smooth response under

stretching the latter is given by [16]

$$F_{\text{str}}(k, f) = \int_{x(k,0)}^{x(k,f)} dx \tilde{f}(k, x) = fx(k, f) - \int_0^f df' x(k, f'), \quad (9)$$

where $\tilde{f}(k, x)$ denotes the inverse function of $x(k, f)$ with respect to f . Accordingly, at a given k , the work for stretching upon increasing the force from f_a to f_b becomes

$$\begin{aligned} - \int_{f_a}^{f_b} df x(k, f) &= F_{\text{str}}(k, f_b) - f_b x(k, f_b) - F_{\text{str}}(k, f_a) + f_a x(k, f_a) \\ &= F_{\text{tot}}(k, f_b) - F_{\text{tot}}(k, f_a). \end{aligned} \quad (10)$$

This just means that the stretching at fixed k is assumed to take place quasi-statically, i.e. the variation of the force is supposed to occur on a time scale much slower than the correlation time of end-to-end distance fluctuations. Inserting this result into EQ. (8) yields

$$W_f(t) = -[F_{\text{tot}}(k, f(t)) - F_{\text{tot}}(k, D(t))] - \sum_{k=1}^{D(t)-1} [F_{\text{tot}}(k, f(t_{k+1})) - F_{\text{tot}}(k, f(t_k))], \quad (11)$$

from which it becomes clear that W_f equals $(-W)$, if $F_{\text{tot}}(k, f(t))$ is identified with the free energy $F_k(t)$ in the extended Kittel model.

As said in the Introduction, we were able to find analytical results for the work distributions, if the backward rates in EQ. (6) were negligible. We are interested here in the model itself and will not make attempts to assign values to the parameters in the transition rates $\lambda(t)$ [EQ. (4)] and λ_b [EQ. (6)], and appearing also in the state free energies in EQ. (1), which are connected to real experiments. Nevertheless, with respect to the value of the analytical results in connection with fluctuation theorems, the question arises whether a neglect of backward rates could be acceptable, at least for certain parameter settings. To check this, we have performed kinetic Monte Carlo simulations of the stochastic process. Because the forward rate from EQ. (4) can be written as $\lambda(t) = \lambda_b G \exp(-\beta\Delta + \beta vt)$ and the free energy differences appearing in EQ. (7) by $[F_k(t_{k+1}) - F_k(t_k)] = -(k-1)v(t_{k+1} - t_k)$, the dynamics and energetics of the model is completely specified by the parameters λ_b , G , v , and T (and N if we consider a complete unfolding). For illustration and discussion of representative results, we here use Δ , Δ/k_B , and ν^{-1} as units for energy, temperature and time, respectively.

Simulations were performed for fixed $N = 10$, $T = 1$ and $\lambda_b = 1$, and a set of v and G values varying in the intervals $v = 0.01 - 3.3$ and $G = 10 - 1000$, respectively. We always started the unfolding from the fully closed state, i.e. $p_k(0) = \delta_{k,1}$ [47]. Probabilities $p_N(t)$ of complete unfolding (occupation of state N) until time t were determined and an unfolding time t_U defined by requiring that at $t = t_U$ the zipper has unfolded with a probability of 99.9%, i.e. $p_N(t_U) = 0.999$ (see Sec. 5). We then considered the work distributions $\rho(w, t_U)$ at $t = t_U$ and the weighted distributions $\exp(-\beta w)\rho(w, t_U)$, corresponding to the integrand in the average $\langle \exp(-\beta w) \rangle = \int dw \exp(-\beta w)\rho(w, t_U)$ as it appears in the Jarzynski equality. It was found that backward rates turn out to have a minor importance when the forward rates are much

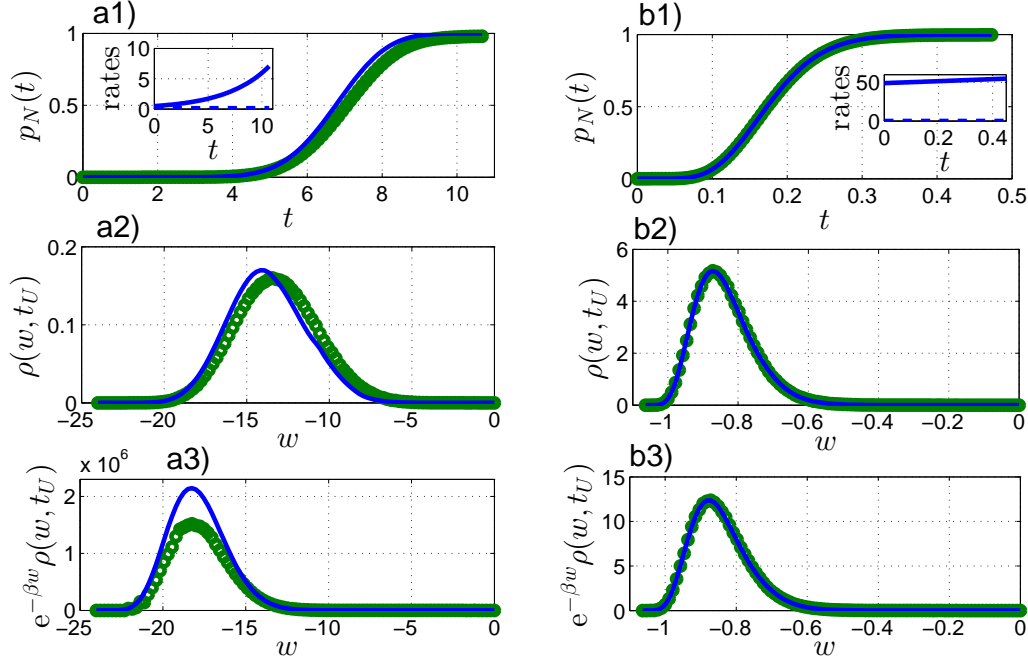


Figure 2. Simulated results of the model presented in SEC. 2 with backward transition rate $\lambda_b = 0.133$ (blue lines) in comparison with analytical results (see Sec. 4) when neglecting backward transitions (green circles). Panels labeled with a) and b) refer to degeneracy factors $G = 10$ and $G = 1000$, respectively. The remaining parameters are $T = 1$, $v = 0.25$, and $N = 10$. Panels a1) and b1) show the probability $p_N(t)$ that the zipper has unfolded up to time t and the insets depict the time-dependence of the forward transition rates (full line); the constant backward rate is indicated by the dashed line. Panels a2) and b2) display the probability densities $\rho(w, t_U)$ at the unfolding time t_U ($t_U = 10.67$ for $G = 10$ and $t_U = 0.47$ for $G = 1000$), and panels a3) and b3) the weighted probability densities $\exp(-\beta w)\rho(w, t_U)$.

larger than the backward rates during the whole unfolding process or when v is large enough. Specifically, the error in calculating $\langle \exp(-\beta w) \rangle$ is smaller than 5% when $v \gtrsim 3$ for $G = 10$, $v \gtrsim 0.1$ for $G = 100$, and $v \gtrsim 0.01$ for $G = 1000$. As representative examples we show in FIG. 2 simulated results (blue lines) with nonzero backward rates in comparison with analytical results (green circles) for zero backward rates (see Sec. 4) for $p_N(t)$, $\rho(w, t_U)$, and $\exp(-\beta w)\rho(w, t_U)$, and parameters $v = 0.25$, $G = 10$ [panels labeled with a)] and $G = 1000$ [panels labeled with b)]. As can be seen from the figure, for the case of large G (small backward transitions), the simulated results for $\exp(-\beta w)\rho(w, t_U)$ are almost indistinguishable from the analytical results.

The remaining part of the paper is organized as follows. In SEC. 3 we specify the equations for the time evolution of the state probabilities [EQ. (12)] and for the quantities describing the energy transformations [EQ. (17)]. In SEC. 4 we derive exact analytical solutions of these equations and in SEC. 5 we discuss our findings.

3. Time evolution of state probabilities and work

Let $p_k(t)$, $k = 1, \dots, N$, be the occupation probabilities of the k -th state. The time evolution of these functions is governed by the Pauli rate equation with transition rates given by EQ. (4). Formally speaking, the unzipping process is described by the time-inhomogeneous Markov process $\mathbf{D}(t)$, where $\mathbf{D}(t) = k$ if the system resides in state k at time t . The Pauli rate equation can be written as

$$\frac{d}{dt}\mathbb{R}(t) = \mathbb{L}(t)\mathbb{R}(t), \quad \mathbb{R}(0) = \mathbb{I}, \quad (12)$$

where \mathbb{I} is the $(N \times N)$ unity matrix, $\mathbb{L}(t)$ is the $(N \times N)$ matrix of the transition rates,

$$\mathbb{L}(t) = \begin{pmatrix} -\lambda(t) & 0 & \dots & \dots & 0 \\ \lambda(t) & -\lambda(t) & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -\lambda(t) & 0 \\ 0 & \dots & 0 & \lambda(t) & 0 \end{pmatrix}, \quad (13)$$

and $\mathbb{R}(t)$ is the $(N \times N)$ matrix of the transition probabilities with the matrix elements

$$R_{ij}(t) = \text{Prob} \{ \mathbf{D}(t) = i \mid \mathbf{D}(0) = j \}. \quad (14)$$

The matrix of transition probabilities evolves an arbitrary column vector of the initial occupation probabilities, $p(t) = \mathbb{R}(t)p(0)$. In the following, we always start with the completely closed zipper, that means $p_k(0) = \delta_{k1}$. The individual occupation probabilities then are

$$p_i(t) = R_{i1}(t). \quad (15)$$

According to EQ. (7) the work $\mathbf{W}(t)$ is a *functional* of the process $\mathbf{D}(t)$. For an analytical treatment it is useful to introduce the *augmented process* $\{\mathbf{W}(t), \mathbf{D}(t)\}$ [49, 50, 33] which describes both the work and the state variable. This augmented process is again a time non-homogeneous Markov process and its one-time properties are described by $(N \times N)$ matrix $\mathbb{G}(w, t)$ with the matrix elements

$$G_{ij}(w, t) = \lim_{\epsilon \rightarrow 0} \frac{\text{Prob} \{ \mathbf{W}(t) \in (w, w + \epsilon) \text{ and } \mathbf{D}(t) = i \mid \mathbf{W}(0) = 0 \text{ and } \mathbf{D}(0) = j \}}{\epsilon}. \quad (16)$$

The time evolution of $\mathbb{G}(w, t)$ is given by [49, 50, 33]

$$\frac{\partial}{\partial t}\mathbb{G}(w, t) = \left[-\frac{\partial}{\partial w} \dot{\mathbb{E}}(t) + \mathbb{L}(t) \right] \mathbb{G}(w, t), \quad \mathbb{G}(w, 0) = \delta(w)\mathbb{I}. \quad (17)$$

Here $\dot{\mathbb{E}}(t)$ is the diagonal matrix $\dot{\mathbb{E}}(t) = \text{diag}\{\dot{E}_1(t), \dots, \dot{E}_N(t)\}$. Notice that $\mathbb{R}(t) = \int_{-\infty}^{\infty} dw \mathbb{G}(w, t)$. EQ. (17) represents a hyperbolic system of N^2 coupled partial differential equations with time-dependent coefficients. Its exact solution will be given in the following SEC. 4.

The matrix $\mathbb{G}(w, t)$ provides a complete description of the energetics of the unzipping process. The joint probability density for the internal energy $\mathbf{U}(t) = E_{\mathbf{D}(t)}(t)$

and the work $W(t)$ performed on the system during the time interval $[0, t]$ (regardless of the final state of the system at the time t) is given by

$$\xi(u, w, t) = \sum_{i=1}^N \delta[u - E_i(t)] G_{i1}(w, t) , \quad (18)$$

where $\delta(x)$ is the Dirac δ -function. The last function already yields the probability density for the work,

$$\rho(w, t) = \int_{-\infty}^{\infty} du \xi(u, w, t) . \quad (19)$$

An analogous integration over the work variable w gives the probability density of $U(t)$. Furthermore, the first law of thermodynamics implies $U(t) - U(0) = U(t) = W(t) + Q(t)$ (note that $U(0) = 0$ for our setting) and, accordingly, $\xi(u, w, t)$ gives also the probability density of the heat $Q(t)$ transferred from the reservoir during the time interval $[0, t]$ [35].

The mean values of the internal energy, work and heat, are then

$$U(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} du dw u \xi(u, w, t) = \sum_{i=1}^N E_i(t) p_i(t) , \quad (20)$$

$$W(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} du dw w \xi(u, w, t) = \sum_{i=1}^N \int_0^t dt \dot{E}_i(t) p_i(t) , \quad (21)$$

$$Q(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} du dw (u - w) \xi(u, w, t) = \sum_{i=1}^N \int_0^t dt E_i(t) \dot{p}_i(t) . \quad (22)$$

Notice that these mean values can be also calculated directly from the solution of the Pauli rate equation (12). However, for higher moments, we already need the function $\xi(u, w, t)$. For example, the variances discussed in SEC. 5 are given by

$$[\Delta U(t)]^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} du dw u^2 \xi(u, w, t) - [U(t)]^2 , \quad (23)$$

$$[\Delta W(t)]^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} du dw w^2 \xi(u, w, t) - [W(t)]^2 , \quad (24)$$

$$[\Delta Q(t)]^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} du dw (u - w)^2 \xi(u, w, t) - [Q(t)]^2 . \quad (25)$$

4. Solution of the model

Owing to the simple two-diagonal structure of the matrix $\mathbb{L}(t)$ in EQ. (13), the system (12) can be solved by simple integrations. By defining

$$\Lambda(t, t') = \int_{t'}^t dt'' \lambda(t'') = \frac{\alpha}{\beta v} [\exp(\beta v t) - \exp(\beta v t')] , \quad (26)$$

where $\alpha = g \exp(-\beta d)$, a recursive treatment of EQ. (12) yields

$$\begin{aligned}
R_{ij}(t) &= 0, & i < j, \\
R_{jj}(t) &= \exp[-\Lambda(t, 0)], & j = 1, \dots, N, \\
R_{ij}(t) &= \int_0^t dt' \exp[-\Lambda(t, t')] \lambda(t') R_{i-1j}(t'), & j < i < N, \\
R_{Nj}(t) &= \int_0^t dt' \lambda(t') R_{N-1j}(t'), & j = 1, \dots, N.
\end{aligned} \tag{27}$$

When solving these recursive relations, we obtain the lower triangular matrix with the nonzero matrix elements

$$R_{j+kj}(t) = \frac{[\Lambda(t, 0)]^k}{k!} \exp[-\Lambda(t, 0)], \quad j = 1, \dots, N-1; \quad k = 0, \dots, N-1-j, \tag{28}$$

$$R_{Nj}(t) = \frac{[\Lambda(t, 0)]^{N-j}}{(N-j)!} \exp[-\Lambda(t, 0)] {}_1F_1(1, N+1-j; \Lambda(t, 0)), \quad j = 1, \dots, N. \tag{29}$$

where ${}_1F_1(a, b; x)$ denotes the confluent hypergeometric function.

For solving EQ. (17) we first perform a Laplace transform with respect to the work variable w . To keep the notation simple, we use the same symbols for the original functions and for the transformed ones. The transformed functions will be distinguished by explicitly giving the complex variable s conjugate to w . After performing the Laplace transformation, we get the system of ordinary differential equations

$$\frac{\partial}{\partial t} \mathbb{G}(s, t) = \left[-s \dot{\mathbb{E}}(t) + \mathbb{L}(t) \right] \mathbb{G}(s, t), \quad \mathbb{G}(s, 0) = \mathbb{I}. \tag{30}$$

The matrix which multiplies $\mathbb{G}(s, t)$ on the right hand side is again a lower two-diagonal one. Therefore, similarly to EQ. (27), we find the recursive relation

$$\begin{aligned}
G_{ij}(s, t) &= 0, & i < j, \\
G_{jj}(s, t) &= \exp[-\Lambda(t, 0)] \exp\{-s[E_j(t) - E_j(0)]\}, & j = 1, \dots, N, \\
G_{ij}(s, t) &= \int_0^t dt' \exp[-\Lambda(t, t')] \exp\{-s[E_i(t) - E_i(t')]\} \lambda(t') G_{i-1j}(s, t'), & j < i < N, \\
G_{Nj}(s, t) &= \int_0^t dt' \exp\{-s[E_N(t) - E_N(t')]\} \lambda(t') G_{N-1j}(s, t'), & j = 1, \dots, N.
\end{aligned} \tag{31}$$

Notice that the matrix $\mathbb{G}(s, t)$ is again a lower triangular one. We now want to solve these recursive relations. It turns out that all matrix elements of the matrix $\mathbb{G}(s, t)$, except the matrix elements $G_{Nj}(w, s)$, $j = 1, \dots, N-1$, can be explicitly evaluated by simple integrations:

$$G_{j+kj}(s, t) = \frac{\left\{ \frac{\alpha}{v(\beta - s)} [\exp(\beta vt) - \exp(svt)] \right\}^k}{k!} \exp[s(j-1)vt] \exp[-\Lambda(t, 0)], \tag{32}$$

for $j = N$, $k = 0$ and also for $j = 1, \dots, N-1$, $k = 0, \dots, N-1-j$. Moreover, we were able to carry out the inverse Laplace transformation of these functions. The resulting diagonal elements $G_{jj}(w, t)$, $j = 1, \dots, N$ are proportional to Dirac δ -functions. The remaining ones possess a finite support, i.e. they are proportional to the differences of the unit-step functions $\Theta(a, b; x) = \Theta(x - a) - \Theta(x - b)$,

$$G_{jj}(w, t) = \delta[w + (j-1)vt] \exp[-\Lambda(t, 0)] , \quad j = 1, \dots, N , \quad (33)$$

$$G_{j+kj}(w, t) = \exp[-\Lambda(t, 0)] \frac{\left(-\frac{\alpha}{v}\right)^k}{(k-1)!} \exp\{\beta[w + (k+j-1)vt]\} \times \\ \times \sum_{l=0}^k \frac{(-1)^l}{l! (k-l)!} \frac{\Theta[(1-j-l)vt, (1-j)vt; w]}{[w + (j+l-1)vt]^{1-k}} . \quad (34)$$

These expressions are valid for $j = 1, \dots, N-1$ and $k = 1, \dots, N-1-j$.

It remains to calculate the matrix elements $G_{Nj}(s, t)$, $j = 1, \dots, N$. The inverse Laplace transformation of the last equation in the recursive scheme (31) is

$$G_{Nj}(w, t) = \int_0^t dt' \lambda(t') G_{N-1j} \{w - [E_N(t) - E_N(t')], t'\} , \quad j = 1, \dots, N-1 . \quad (35)$$

We insert herein the explicit forms of EQS. (33) and (34). After some algebra we finally obtain

$$G_{NN-1}(w, t) = \frac{\alpha}{v} \exp\{\beta[w + (N-1)vt]\} \exp\left\{-\Lambda\left[\frac{w}{v} + (N-1)t, 0\right]\right\} \times \\ \times \Theta[-(N-1)vt, -(N-2)vt; w] , \quad (36)$$

$$G_{Nj}(w, t) = \frac{\alpha \left(-\frac{\alpha}{v}\right)^{N-1-j}}{(N-2-j)!} \exp\{\beta[w + (N-1)vt]\} \sum_{l=0}^{N-1-j} \frac{(-1)^l}{l! (N-1-j-l)!} \times \\ \times \left\{ F_{jl} \left[\frac{w + (N-1)vt}{v(N-j)}, \frac{w + (N-1)vt}{v(N-j-l)}; w, t \right] \Theta[-(N-1)vt, -(j+l-1)vt; w] + \right. \\ \left. + F_{jl} \left[\frac{w + (N-1)vt}{v(N-j)}, t; w, t \right] \Theta[-(j+l-1)vt, -(j-1)vt; w] \right\} , \quad (37)$$

Here $j \in \{1, \dots, N-2\}$ and we have introduced the abbreviation

$$F_{jl}(a, b; w, t) = \int_a^b dt' \exp[-\Lambda(t', 0)] [w + (N-1)vt - (N-j-l)vt']^{N-j-2} . \quad (38)$$

The main results of this Section are EQS. (28) and (29), which give the solution of the Pauli rate equation (12), and EQS. (33)-(38) which present the solution of EQ. (17). We now turn to the discussion of these results.

5. Discussion

In Kittel's work [8] the *equilibrium* properties of the zipper are studied. The mean number of open links in equilibrium always increases with increasing temperature, but the form of this increase is different for the degeneracy factor $G = 1$ (cf. FIG. 3a1)) and for $G > 1$ (cf. FIG. 3a2)). For $G = 1$, the mean number of open links increases smoothly with a concave curvature, while for $G > 1$, the curve exhibits a sharp increase in a narrow temperature interval and resembles a first-order phase transition. In the following discussion of representative results for the nonequilibrium dynamics and energetics, we use d , d/k_B , and ν^{-1} as units for energy, temperature and time, respectively.

In the unidirectional unzipping process, the time evolution towards the completely unzipped state is again sensitive to the degeneracy factor G . Let us define an unfolding time t_U by the condition that at $t = t_U$ the zipper has completely unfolded with a probability of 99,9%, i.e.

$$p_N(t_U) = R_{N1}(t_U) = 1 - \epsilon, \quad (39)$$

where $\epsilon = 0.001$. With respect to the N dependence of t_U (and other quantities to be discussed below), we found that its behavior is similar for $N = 2$ and $N > 2$, and we therefore restrict the following discussion to the two-state case $N = 2$. EQ. (39) can then be inverted after inserting $R_{N1}(t_U)$ from EQ. (29) in EQ. (39), yielding

$$t_U = [1/(\beta v)] \ln [1 - (\beta v/\alpha) \ln \epsilon]. \quad (40)$$

For small temperatures [large argument of the exponential and/or small prefactor g in EQ. (4)] the transitions are driven predominantly by the magnitude of the energy gap between the closed and the opened state. For large temperatures [small argument of the exponential and/or large prefactor g in EQ. (4)] by contrast, the dynamics is governed by the entropy difference between the closed and the opened state, and hence by the degeneracy factor G . The dependence of the unfolding time t_U on G and T is plotted in FIG. 3b) for a representative set of parameters. For any fixed nonzero temperature, the unfolding time decreases with increasing G , while for fixed G , the temperature-dependence of the unfolding time exhibits a maximum at a temperature $T = T_{\max}(G)$, where $T_{\max}(G)$ decreases with increasing G , see FIG. 3b).

Let us now consider a certain temperature T_0 and call, for this temperature, the *fast-unzipping regime* and *slow-unzipping regime* the ranges of G -values, where $T_{\max}(G) < T_0$ and $T_{\max}(G) > T_0$, respectively. In these two regimes the dynamics and energetics of the molecular zipper exhibit a qualitatively different behavior. In particular we find (i) a different time-dependence of the N -th state's occupation probability, cf. FIGS. 4b) and 5b), (ii) a different form of the curves describing the work needed to open the zipper, cf. FIGS. 4c) and 5c), (iii) a different mean value of heat accepted by the zipper during the unzipping, cf. FIGS. 6a1) and 6b1), and iv) different values of the variances of the internal energy and heat during the unzipping, cf. FIGS. 6a2) and 6b2). These features will be now discussed in more detail.

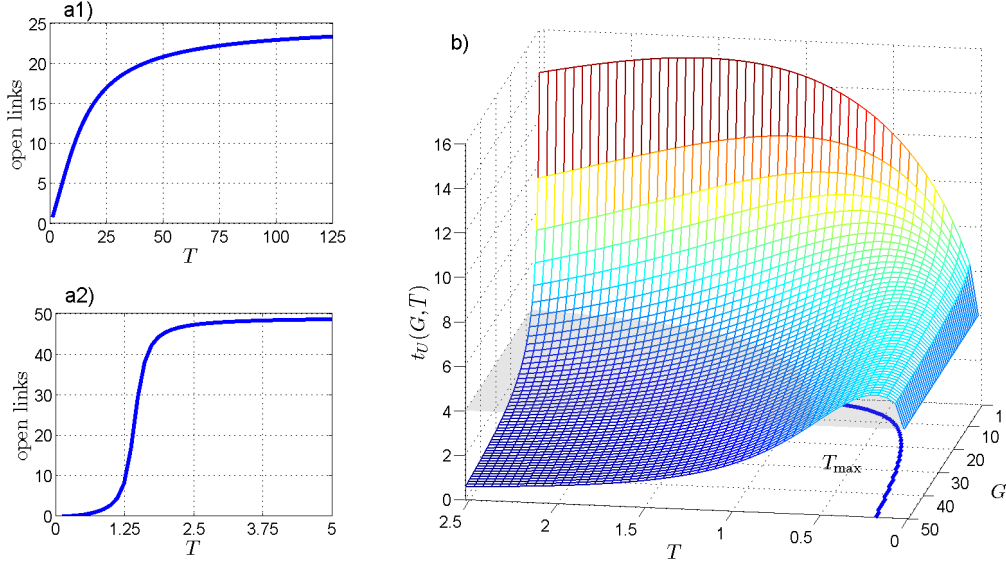


Figure 3. Left two panels: Mean number of open links in equilibrium for Kittel’s molecular zipper as a function of the reservoir temperature T for $N = 50$ links, and (a1)) $G = 1$ and (a2)) $G = 2$. Right panel b): Unfolding time t_U , EQ. (40), as a function of T and G for $T_0 = 7.5$, $v = 0.25$, and $N = 2$. Above (below) the horizontal plane in the graph, the opened (closed) state is energetically favored. In the base plane, the temperature T_{\max} of maximal unfolding time in dependence of the degeneracy factor G is shown.

FIG. 4 illustrates the slow-unzipping regime. The probability $p_N(t)$ that the zipper has reached the opened state until time t first increases slowly. After the time d/v the energy of the opened state becomes lower than that of the closed one. This leads to more frequent transitions and accordingly $p_N(t)$ increases more rapidly. Notice that the curves exhibit a change of their second derivative. The work probability density (WPD) during the unzipping has a maximum located inside its finite support, cf. FIG. 4c). The value of the work at the left (right) border of the support equals the work done on the zipper when it dwells during the time interval $[0, t_U]$ in the opened (closed) state. From the position of the WPD peak we can conclude that, for a typical trajectory of the stochastic process $D(t)$, the work consists of two comparable fractions. The first (second) part of the work is performed while the system dwells in the closed (opened) state. At time t_U , 0.1% of the trajectories will give molecules still residing in the zipped state. These trajectories contribute to the singular (δ -function) components of WPDs, which are depicted in FIG. 4c) by the vertical arrows [33, 32, 34, 35]. Note that the curves plotted for the temperatures $T = 1$ and $T = 1.25$, which are close to the temperature $T_{\max}(G)$ for $G = 1$, cf. FIG. 3b), become similar to the curves for these temperatures obtained in the fast-unzipping regime for $G = 50$, see FIG. 5.

FIG. 5 illustrates the fast-unzipping regime. The probability $p_N(t)$ rapidly increases from the very beginning of the process, cf. FIG. 5b), i.e. the zipper opens before the

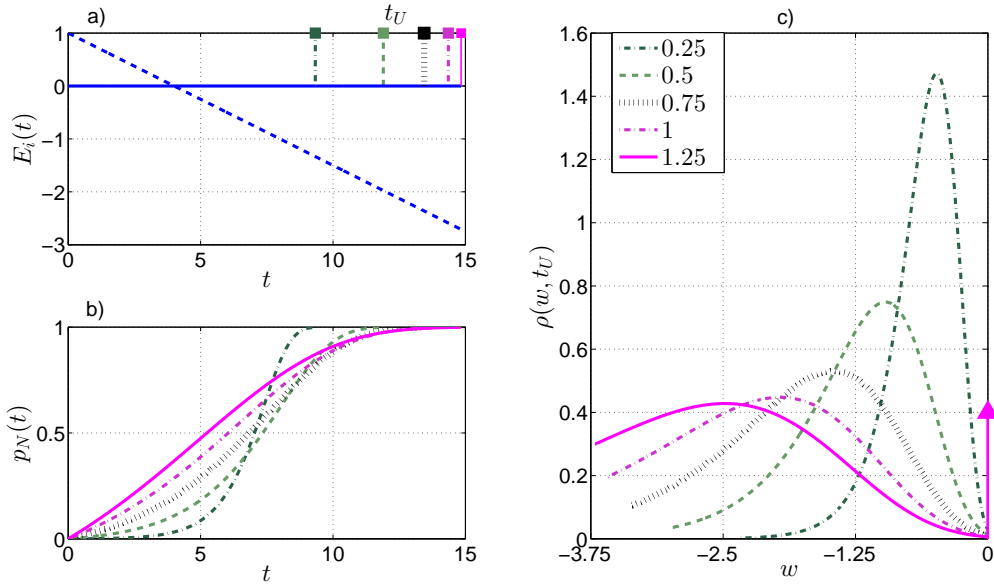


Figure 4. Dynamics of a) the energy levels [EQ. (1)], b) the occupation probability $p_N(t)$ of the opened state [EQ. (15)], and c) the probability density $\rho(w, t_U)$ of the work [EQ. (19)] for $G = 1$ and several values of the reservoir temperature T . The other parameters are the same as in FIG. 3. The assignment of the line styles to the temperature given in the legend of c) applies also to parts a) and b). In a) we compare the dynamics of the two energy levels with the unfolding time t_U [EQ. (40)] at different temperatures marked by the vertical lines. The full (dashed) line gives the energy $E_1(t)$ ($E_2(t)$) of the closed (unzipped) state. The arrows in c) represent the weights and the positions of the δ -functions, which form the singular part of the probability density.

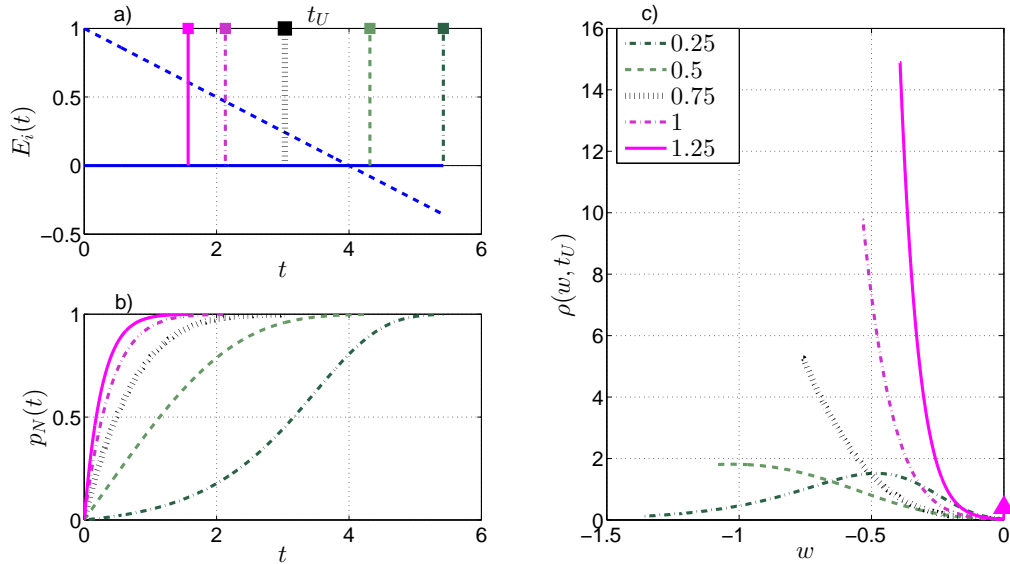


Figure 5. Same quantities as in FIG. 4 for $G = 50$ and otherwise the same set of parameters.

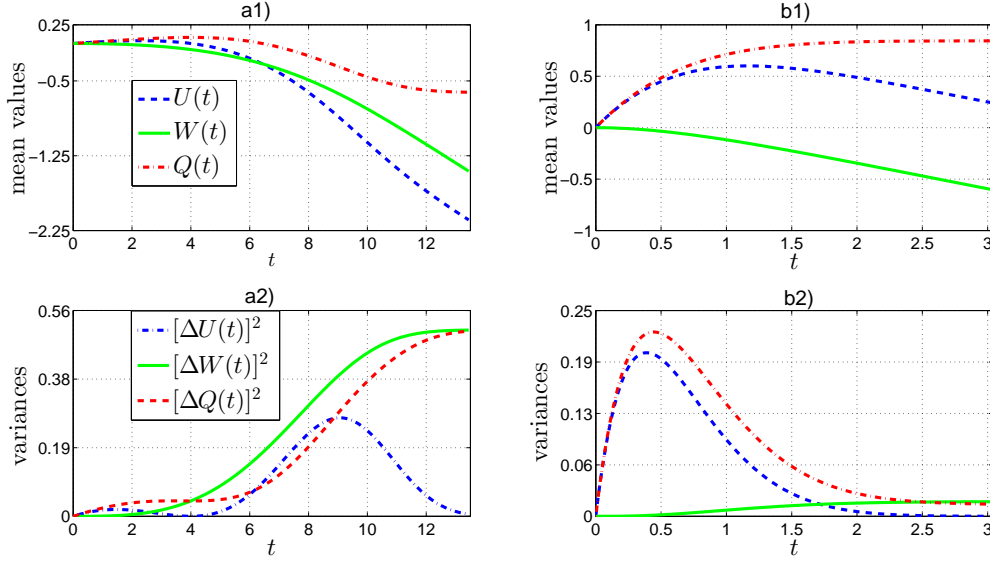


Figure 6. Upper two panels: Mean value of the internal energy [Eq. (20)], work [Eq. (21)] and heat [Eq. (22)] as a function of time for a1) $G = 1$ and b1) $G = 50$. Lower two panels: Variances of internal energy [Eq. (23)], work [Eq. (24)] and heat [Eq. (25)] as a function of time for a2) $G = 1$ and b2) $G = 50$. The temperature is $T = 0.75$ and the other parameters are the same as in FIG. 3.

opened state becomes energetically preferred. The maximum of the WPD is located at the left border of its support, cf. FIG. 5c). This means that, for the majority of the trajectories, the substantial part of the work is done while the system dwells in the opened state. Note that the curves plotted for the temperature $T = 0.25$, which is close to the boundary temperature $T_{\max}(G)$ for $G = 50$, cf. FIG. 3b), become similar to the curves for this temperature in the fast-unzipping regime for $G = 1$, see FIG. 4.

FIG. 6 illustrates the dynamics of the thermodynamic quantities (20)–(25) in the two unzipping regimes. For an arbitrary trajectory of $D(t)$ which resides during the time interval $[t', t]$ in the i -th state, the work performed on the system is $E_i(t) - E_i(t')$, cf. EQ. (7). In our model, the energies of the states decrease linearly with time and accordingly the mean work is a monotonically decreasing function of time, see FIGS. 6a1) and 6b1). Heat is exchanged with the reservoir when the molecule changes its state. It is absorbed by the molecule if the transition brings the molecular zipper to a state with higher energy. Since in our setting the transitions are unidirectional, the molecule necessarily absorbs heat up to the time $t_E = d/v$, where the energies of the states become the same, cf. FIGS. 4a) and 5a). For times $t > t_E$, the molecule delivers heat to the environment. In the slow-unzipping regime we have $t_U > t_E$. This implies that the mean heat first increases and then decreases, cf. FIG. 6a1). By contrast, in the fast-unzipping regime where $t_U < t_E$, the mean heat monotonically increases, cf. FIG. 6b1). Finally, due to the transitions to the state with higher energy at the very beginning of the process, cf. FIGS. 4 and 5, the mean internal energy (20) develops a single maximum.

The variances of $U(t)$, $W(t)$ and $Q(t)$, cf. EQS. (23)–(25), are plotted in FIGS. 6a2) and FIGS. 6b2). If the variance of the internal energy approaches zero, the variance of the work becomes equal to that of the heat. All variances approach constant values at large times. In fact, for $t > t_U$, almost all trajectories have brought the molecule into the opened state. As a consequence, the increments to the work, heat and internal energy are nearly constant during the time interval $[t_U, t]$. Therefore the form of their probability densities does not change, the curves just move along the energy axis.

The probability density for the internal energy has no continuous component. In general it consists of N δ -functions located at the energies of the individual states. The corresponding weights are given by the occupation probabilities of the states. In the slow-unzipping regime, $[\Delta U(t)]^2$ vanishes at time $t = 0$ and at time $t_E = d/v$, when the state energies are equal, and it becomes very close to zero for $t \gtrsim t_U$. In between these time instants, the variance develops a maximum, cf. FIG. 6a2). In the fast-unzipping regime, the function $[\Delta U(t)]^2$ displays just one maximum, cf. FIG. 6b2), because $t_U < t_E$.

In both the fast and slow unzipping regime, the variance of the work monotonically increases and approaches a constant for large times, cf. FIG. 6a2) and FIG. 6b2). The absolute value of the first derivate of the variance $[\Delta Q(t)]^2$ is given by the product of the two quantities, the energy difference between the closed and the opened states at time t and the probability that the zipper opens at the time t , which were shown in FIGS. 4 and 5. In the slow-unzipping regime, the majority of the transitions occurs during the time interval $[t_E, t_U]$. During this time interval, the energy difference between the states monotonically increases. In the fast-unzipping regime, by contrast, nearly all transitions take place before time $t_E/2$. The sooner the transition, the larger is the amount of transferred heat. As a result, at small times, the heat probability density has one peak at a large value, coming from the trajectories with a transition to the opened state, and another peak at zero heat exchange, originating from trajectories without a transition. With increasing time the amount of the trajectories with a transition to the opened state increases rapidly, and the peak close to zero heat moves towards the peak at a large heat value. This explains the behavior of the heat variances in FIGS. 6a2) and 6b2).

Finally, one may ask how our results are affected if the variation (“static disorder”) in base pairing energies is included in the modeling. To this end we have performed Monte Carlo simulations [48] of the stochastic process for a molecular zipper with $N = 10$ states, where the initial energies $E_k(0)$, $k = 1, \dots, N$, in EQ. (1) are given by $E_k(0) = (k-1)(\Delta + \eta_k)$, corresponding to different losses of energies due to variations in base pair bondings. The η_k were chosen as random numbers from a box distribution in the interval $[-\Delta/3, \Delta/3]$. Results from these simulations for a number of realizations of this disorder were compared to the predictions of the analytical theory for the “ordered case”, where $E_k(0) = (k-1)\Delta$. We found that, for typical realizations of sets of η_k , the shapes of the work distributions are very similar, while the peak positions and peak heights are shifted slightly. Also the probability distributions $p_N(t)$ for the zipper to

fully open until time t are, for these sets of η_k , very similar in shape. Analogous to the peak positions of the work distributions, the onset of opening shifts slightly from realization to realization. The shifts of the peaks and of the onset of the opening are controlled by the largest base pair bonding (largest η_k), which governs the unfolding time.

Acknowledgments

Support of this work by the Ministry of Education of the Czech Republic (project No. MSM 0021620835), by the Grant Agency of the Charles University (grant No. 143610, and grant No. 301311), by the Charles University in Prague (project No. SVV-2012-265 301), and by the Deutsche Akademische Austauschdienst (DAAD, project No. MEB101104) is gratefully acknowledged.

References

- [1] Poland D and Scheraga H, 1966 *J. Chem. Phys.* **45** 1464
- [2] Hanke A and Metzler R, 2003 *J. Phys. A: Math. Gen.* **36** L473
- [3] Bicout D J and Kats E, 2004 *Phys. Rev. E* **70** 010902(R)
- [4] Fogedby H C and Metzler R, 2007 *Phys. Rev. Lett.* **98** 070601
- [5] Metzler R, Ambjörnsson T, Hanke A and Fogedby H C, 2009 *J. Phys.: Condens. Matter* **21** 034111
- [6] Gibbs J H and Dimarzio E A, 1959 *J. Chem. Phys.* **30** 271
- [7] Crothers D M, Kallenbach N R and Zimm B H, 1965 *J. Mol. Biol.* **11** 802
- [8] Kittel C, 1969 *Am. J. Phys.* **37** 917
- [9] Liphardt D, Onoa B, Smith S, Tinoco I, and Bustamante C, 2001 *Science* **292** 733
- [10] Liphardt D, Dumont S, Smith S, Tinoco I, and Bustamante C, 2002 *Science* **296** 1832
- [11] Lang M J and Block S M, 2003 *Am. J. Phys.* **71** 201
- [12] Onoa B, Dumont S, Liphardt J, Smith S, Tinoco I Jr. and Bustamante C, 2003 *Science* **299** 1892
- [13] Ritort F, 2006 *J. Phys.: Condens. Matter* **18** R531
- [14] Ritort F, 2008 *Adv. Chem. Phys.* **137** 31
- [15] Tinoco I Jr and Bustamante C, 1999 *J. Mol. Biol.* **293** 271
- [16] Manosas M and Ritort F, 2005 *Biophys. J.* **88** 3224
- [17] Thirumalai D and Hyeon C, 2005 *Biochemistry* **44** 4957
- [18] Finkelstein A V, 2004 *Proteins: structural, thermodynamic and kinetic aspects. In J. L. Barrat and J. Kurchan, editors, Slow relaxations and nonequilibrium dynamics* (Berlin: Springer-Verlag)
- [19] R. Green, and H. F. Noller 1997, *Annu. Rev. Biochem.* **66** 679
- [20] Ramakrishnan V, 2002 *Cell* **108** 557
- [21] Kornberg A and Baker T A, 1992 *DNA Replication* (New York: W. H. Freeman and Company)
- [22] Tackett A J, Morris P D, Dennis R, Goodwin T E, and Raney K D, 2001 *Biochemistry* **40** 543
- [23] Palassini M and Ritort F, 2011, *Phys. Rev. Lett.* **107** 060601
- [24] Esposito M, Van den Broeck C, 2010 *Phys. Rev. Lett.*, **104** (2010) 090601; *ibid.*, *Phys. Rev. E* **82** 011143
- [25] Van den Broeck C, Esposito M, 2010 *Phys. Rev. E* **82** 011144
- [26] Seifert U, 2008 *Eur. Phys. J. B* **64** 423
- [27] Braun O, Hanke A, and Seifert U, 2004 *Phys. Rev. Lett.* **93** 158105
- [28] Crooks G E, 1999 *Phys. Rev. E* **60** 2721
- [29] Jarzynski C, 1997, *Phys. Rev. Lett.* **78**, 2690.
- [30] Chatelain C and Karevski D, 2006 *J. Stat. Mech.* P06005

- [31] Híjar H, Quintana J and Sutmann G, 2007 *J. Stat. Mech.* P04010
- [32] Chvosta P, Reineker P and Schulz M, 2007 *Phys. Rev. E* **75** 041124
- [33] Subrt E and Chvosta P, 2007 *J. Stat. Mech.*, P09019.
- [34] Einax M and Maass P, 2009 *Phys. Rev. E* **80** 020102
- [35] Chvosta P, Einax M, Holubec V, Ryabov A and Maass P, 2010 *J. Stat. Mech.* P03002
- [36] Mazonka O and Jarzynski C, 1999 *arXiv:cond-mat/9912121v1*
- [37] Baule A and Cohen E D G, 2009 *Phys. Rev. E* **80** 011110
- [38] Manosas M, Mossa A, Forns N, Huguet J M and Ritort F, 2009 *J. Stat. Mech.* P02061
- [39] Lubensky D K and Nelson D R, 2000 *Phys. Rev. Lett.* **85** 1572
- [40] Engel S, Alemany A, Forns N, Maass P and Ritort F, 2011 *Phil. Mag. B* **91** 2049
- [41] Mossa A, Manosas M, Forns N, Huguet J M and Ritort F, 2009 *J. Stat. Mech.* P02060
- [42] Bell I G, 1978 *Science* **200** 618
- [43] Smaller (larger) γ would correspond to saddle points lying closer to state k ($k+1$) in configuration space. The derivations in Secs. 3 and 4 can be performed analogously for any γ and F_b .
- [44] Cocco S, Monasson R and Marko J F, 2001 *Proc. Nat. Acad. Sci.* **98** 8608
- [45] Alemany A, Ribezzi M and Ritort F, 2011 *AIP Conference Proceedings* **1332** 96
- [46] Force controlled experiments are achievable with magnetic tweezers or with optical tweezers operating in the force clamp mode. In the latter case, the conjugate variable to the applied force f should better be taken as the trap-pipette distance (rather than the molecular extension) because for this definition the fluctuation theorems are obeyed. Both definitions differ only by a boundary term involving the final and initial forces f_f and f_i , respectively (for details, see [45]).
- [47] In this case one can show that $\langle \exp(-\beta w) \rangle = \phi(t) \exp\{-\beta[F(t) - F(1, 0)]\} = \phi(t) \exp[-\beta F(t)]$, where $F(1, 0) = 0$ from Eq. (1) and $F(t) = -\beta^{-1} \ln[(A^N - 1)/(A - 1)]$, $A = \exp[-\beta(\Delta + vt)]/G$, is the free energy for an equilibrated system with state energies $F(k, t)$ (protocol variables) at time t ; $\phi(t)$ is the probability for the zipper to be in the fully closed state $k = 1$ under the reversed protocol, if the states are initially distributed according to the equilibrium distribution $\exp[F(t) - F(k, t)]$.
- [48] Holubec V, Chvosta P, Einax M and Maass P, 2011 *Europhys. Lett.* **93** 40003
- [49] Imparato A and Peliti L, 2005 *Europhys. Lett.* **69** 643
- [50] Imparato A and Peliti L, 2005 *Europhys. Lett.* **70** 740